Faculty of Resource Science and Technology

A Meta-analysis of Gene Expression of Biomarkers and Construction of Multigene Prognosis Assessment Model in Nasopharyngeal Carcinoma

Sim Chor Chien

Master of Science
2020

A Meta-analysis of Gene Expression of Biomarkers and Construction of Multigene Prognosis Assessment Model in Nasopharyngeal Carcinoma

Sim Chor Chien

A thesis submitted

In fulfilment of the requirements of the degree of Master of Science

(Medical Genetics)

Faculty of Resource Science and Technology
UNIVERSITI MALAYSIA SARAWAK
2020

DECLARATION

I, Sim Chor Chien (17020107), from Faculty of Resource Science and Technology, hereby declare that the work entitled 'A Meta-analysis of Gene Expression of Biomarkers and Construction of Multigene Prognosis Assessment Model in Nasopharyngeal Carcinoma' is my original work. I have not copied from other students' work or from any other sources except where due references acknowledgement is made explicitly in the text, nor has any part been written for me by another person. The thesis has not been accepted for any degree and is not concurrently submitted in candidature for any other degree.

_____

Sim Chor Chien

Matric No: 17020107

Date: 4 June 2020

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere appreciation to my motivating supervisor, Prof. Dr Edmund Sim for his guidance, supervision and valuable suggestion and insightful comments that enable me to acquire better understanding. His helping hands and expertize always help me in producing better quality work. Thank you, Prof. E.

A special thanks to my co-supervisor, Dr Chong Yee Ling for giving me the direction in doing this research. Her professional guidance has provided me with intelligent assistance.

Besides, I would like to thank my dear colleagues, Ms Ng Kher Lee, Ms Felicia Kavita, Ms Cassandra Chee, Ms Yew Keh Li and Ms Cassandra Chong for the support and accompany. A special thanks also for my friend, Yee Siew who has provided me with much IT support throughout the research. Thank you for easing this stressful journey!

I place on record, my utmost gratitude to my dearest family for the unceasing love, encouragement, support and care. And thank you Stewart, for the constant prayers and unconditional support. My deepest heart-felt thanks to you all for accompanying me throughout this journey and sharing my stress and tears. Thank you with lots of love.

Last but not least, I would like to thank God the Almighty for giving me the strength, wisdom, faith and boundless grace to pursue this study. Thank You for leading me through ups and downs with Your out-strengthened arms. Thank you, Lord!

ABSTRACT

Nasopharyngeal carcinoma (NPC) is a common cancer in Southeast Asia, including Malaysia. Overall survival rate of NPC patients is poor due to late diagnosis at the advanced stage. Thus, definitive prognostic biomarkers and multigene prognosis assessment models are crucial to better manage NPC. Cyclooxygenase-2 (COX-2) was identified as a significant prognostic biomarker via meta-analysis in head and neck (HNC) cancer, oral squamous cell carcinoma (OSCC) and breast cancer. Prognostic significance of COX-2 in NPC regarding lymph node metastasis has been done via meta-analysis. However, prognostic significance of COX-2 in NPC in term of overall survival (OS) rate analysed via meta-analysis remains unexplored. This study aims to evaluate COX-2 expression with OS and treatment response via meta-analysis by referring to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist. Results show that over expression of COX-2 predicts worse survival rate outcome for NPC (hazard ratio greater than one) but not treatment response (odd ratio less than one). Hazard ratio is a time dependent statistic measure while odd ratio is a time independent statistic measure. Hence, evaluation of COX-2 expression in NPC patients is important for management and treatment of NPC. Most existing NPC prognosis assessment models are constructed with small sample sizes and thus less relevant and applicable for medical practices. This study therefore aimed to construct a multigene NPC prognosis assessment model by pooling results of more significant meta-analysis studies on prognostic biomarkers of NPC via modification of PRISMA. Through this study, a multigene NPC prognosis assessment model consisting of ten prognostic biomarkers, including COX-2 done from this study is constructed. OS of NPC patients can be determined computationally as pooled hazard ratios (HRs) via this model by selecting the over expressed biomarker(s). The model can also predict the NPC prognosis assessment risk score of the

patient too. Through this model, patients can estimate the overall survival rate time effectively. To provide understanding and explanation to the multigene NPC prognosis assessment model generated computationally, pathways and interactions involved by all the prognostic biomarkers are studied to determine the physiological interactions among all the prognostic biomarkers incorporated. The information on the pathways and interactions are extracted from the literature and summarized as a molecular pathway network. The pathway network model explains the possible complex interplay among different prognostic biomarkers in the oncogenesis of NPC.

Keywords: Nasopharyngeal carcinoma, meta-analysis, cyclooxygenase-2, multigene NPC prognosis assessment model, NPC pathway network model.

Meta-analisis Mengenai Ekspresi Gen Faktor Ramalan dan Penjanaan Model Penilaian Prognosis Pelbagai Gen dalam Karsinoma Nasofarinks

ABSTRAK

Karsinoma nasofarinks (NPC) ialah sejenis kanser yang lazim di Asia Tenggara, termasuk Malaysia. Kadar kelangsungan hidup pada pesakit NPC adalah rendah disebabkan oleh diagnosis lambat pada peringkat akhir. Oleh itu, faktor ramalan NPC yang muktamad dan modal penilaian prognosis pelbagai gen adalah penting untuk mengendali NPC dengan lebih baik. Cyclooxygenase-2 (COX-2) dikenali sebagai faktor ramalan yang penting melalui meta-analisis bagi kanser kepala dan leher, karsinoma sel skuamosa mulut dan kanser dada. Kepentingan COX-2 dalam prognosis NPC mengenai metastasis nodus limfa telah ditentukan melalui meta-analisis. Walaubagaimanapun, fungsi COX-2 sebagai faktor ramalan bagi menentukan kelangsungan hidup pada pesakit NPC belum lagi dikaji melalui meta-analisis. Matlamat kajian ini adalah untuk menilai hubungan antara ekspresi COX-2 dengan kadar kelangsungan hidup dan tindak balas rawatan melalui meta-analisis dengan rujukan 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)'. Keputusan menunjukkan bahawa ekspresi COX-2 yang tinggi menjangkakan kadar kelangsungan hidup yang rendah (nisbah bahaya melebihi satu) tetapi tidak berkaitan dengan tindak balas rawatan (nisbah ganjil kurang daripada satu). Nisbah bahaya merupakan pengiraan statistik bergantung pada masa sementara nisbah ganjil merupakan pengiraan statistik yang tidak bergantung pada masa. Jadi, kajian untuk menilai ekspresi COX-2 adalah penting untuk pengurusan dan rawatan NPC yang lebih cekap. Kebanyakan model penilaian prognosis NPC yang sedang ada dijana dengan saiz sampel yang kecil dan menyebabkan model tersebut kurang berkaitan dan berguna dalam bidang perubatan. Oleh itu, kajian tersebut bertujuan untuk menjana model penilaian prognosis NPC pelbagai gen

yang berkomputasi dengan mengumpulkan keputusan meta-analisa mengenai faktor ramalan NPC yang lebih bermakna melalui pengubahsuaian PRISMA. Dengan kaedah tersebut, model penilaian prognosis NPC pelbagai gen yang terdiri daripada sepuluh faktor ramalan, termasuk COX-2 yang dikaji dalam kajian ini telah dijana. Kadar kelangsungan hidup bagi pesakit NPC dapat ditentukan secara berkomputasi sebagai nisbah bahaya keseluruhan (HRs) melalui model ini dengan memilih faktor ramalan yang diekspres lebih. Model tersebut juga dapat meramalkan angka risiko bagi penilaian prognosis NPC. Dengan menggunakan model ini, pesakit NPC dapat menjangka kadar hidup keseluruhan dalam masa yang singkat. Bagi memberikan pemahaman dan penjelasan mengenai model penilaian prognosis NPC pelbagai gen tersebut yang telah dijana secara berkomputasi, laluan dan interaksi antara faktor ramalan yang terlibat telah dikaji dan interaksi fisiologi telah dijelaskan. Maklumat mengenai laluan dan interaksi telah dikumpulkan dari kajian literatur dan diringkaskan dalam satu model rangkaian laluan NPC. Rangkaian tersebut menjelaskan interaksi kompleks antara faktor ramalan dalam onkogenesis NPC.

Kata kunci: *Karsinoma nasofarinks, meta-analisis, cyclooxygenase-2, model penilaian NPC pelbagai gen, model rangkaian laluan NPC.*

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

5-FU	5-fluorouracil
ASR	Age-adjusted rate
BAPN	β-aminopropionitrile
BRD7	Bromodomain-containing protein 7
$\chi 2$	Chi-squared
CCRT	Concurrent chemoradiotherapy
CIs	Confidence intervals
CRP	C-reactive protein
CRT	Chemoradiotherapy
CTAR 1 and 2	C-terminal activating regions 1 and 2
CXCL 5, 9 and 10	CXC motif chemokine ligand 5, 9 and 10
COX-2	Cyclooxygenase-2
DNA	Deoxyribonucleic acid
EBNA	EBV-determined nuclear antigens
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EBV	Epstein-Barr virus
ERCC1	Excision repair cross-complementation group 1
ECM	Extracellular matrix
FN1	Fibronectin 1
Gy	Gray
HRs	Hazard ratios

HNSCC	Head and neck squamous cell carcinoma
HIF-1α	Hypoxia-inducible factor-1 alpha
IC	Induction chemotherapy
ICRT	Induction chemoradiotherapy,
IHC	Immunohistochemistry
IL-6	Interleukin-6
ISH	RNA-in situ hybridization
I^2	I-squared
LDH	Lactate dehydrogenase
LTF	Lactotransferrin
LMP1	Latent membrane protein 1
logHR	Log hazard ratio
LOH	Loss of heterozygosity
LOXL1–4	LOX-like proteins 1-4
MMP9	Matrix metalloproteinase 9
MTA1	Metastasis-associated protein 1
MAP	Mitogen-activated protein
NPC	Nasopharyngeal carcinoma
NGX6	Nasopharyngeal carcinoma associated gene 6
Nor1	Neuron-derived orphan receptor 1
NSAIDs	Non-steroidal anti-inflammatory drugs
NR	Not related
NF-κb	Nuclear factor κ-B
OR	Odd ratio

OSCC	Oral squamous cell carcinoma
OS	Overall survival
PI3K	Phosphoinositol-3-kinase
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PD-L1/ PD-1	Programmed cell death ligand-1/ Programmed cell death-1
PTGS2	Prostaglandin-endoperoxide synthase 2
PGE	Prostaglandins
QUOROM	Quality Of Reporting of Meta-analyses
RTCs	Randomized controlled trials
RR	Risk ratio
RT	Radiotherapy
RT-PCR	Real-time polymerase chain reaction
SPLUNC1	Short palate, lung, and nasal epithelium clone 1
SNP	Single nucleotide polymorphism
SMD	Standardized mean difference
TPZ	Tirapazamine
TGF-α	Transforming growth factor alpha
TSGs	Tumor suppressor genes
UBAP1	Ubiquitin Associated Protein 1
Var	Variance
VEGF	Vascular endothelial growth factor
WHO	World Health Organization
Zn^{2+}	Zinc

CHAPTER 1

INTRODUCTION

Nasopharyngeal carcinoma (NPC) is the formation of malignancy in the tissue of the upper section of pharynx behind the nose (He et al., 2016), more precisely detected at the fossa of Rosenmüller (Lee, Laufer, Ove, Foote & Bonner, 2012). NPC is rare in the Western hemisphere but widely distributed in Southeast Asia countries and in China (Chang & Adami, 2006). In Penisular Malaysia, NPC was utmost by the Chinese, followed by the Malays and the Indians (Prasad & Rampal, 1992). In Sarawak, the Bidayuh indigenous group showed the highest incident rate with age-adjusted rate (ASR) of 31.5, which was 50% higher than that in Hong Kong (Devi, Pisani, Tang & Parkin, 2004). Throughout Malaysia, it is reporting around 4000 new cases every year (Azizah, Nor-Saleha, Noor-Hashimah, Asmah & Mastulu, 2015). The concurrent chemoradiotherapy (CCRT) with or without neoadjuvant/adjuvant cisplatin-based chemotherapy are the current improved standard treatment protocol administrated to locally advanced NPC patients (Fendri et al., 2009; Wu et al., 2014). The major causes of death among NPC patient are due to local recurrence and distant metastasis, because a large number of patients reached advanced stage and/or metastasis at the time of diagnosis (You et al., 2016). Thus, prognosis is important to better increase the survival rate of NPC patients.

A number prognostic biomarkers such as excision repair cross-complementation group 1 (ERCC1) (Yang et al., 2019), latent membrane protein 1 (LMP1) (Chen et al., 2015), vascular endothelial growth factor (VEGF) (Wang et al., 2018), epidermal growth factor receptor (EGFR) (Sun et al., 2013), C-reactive protein (CRP) (Fang, Xu, Wu, Zhang, Li et

al., 2017), hypoxia-inducible factor-1 alpha (HIF-1α) (Xie et al., 2018), lactate dehydrogenase (LDH) (Zhai, Gu, Zhai, Wang & Zhang, 2017), lysyl oxidase (LOX) (Zhang, Wu, Ye & Xue, 2018) and matrix metalloproteinase 9 (MMP9) (Liao, Huang, Zhu, Li, Huang & He, 2016) were identified as significant prognostic indicators to determine overall survival (OS) of NPC patients. Comprehensive studies on these biomarkers through meta-analysis have been carried out to draw a more definitive conclusion for the above biomarkers too. Over expression of cyclooxygenase-2 (COX-2) was frequently identified as a prognostic biomarker in different cancers such as breast cancer (Jana et al., 2014), colorectal cancer (Peng, Zhou, Wang, Mou & Zhao, 2013) and pancreatic cancer (Juuti, Louhimo, Nordling, Ristimaki & Haglund, 2006). Studies of COX-2 as prognostic biomarker in NPC have also been done (Pan et al., 2008; Loong et al., 2009). However, the sample sizes of these studies are often small (n=35 and n=58) and the results are conflicting. Even though meta-analysis of COX-2 in NPC regarding lymph node metastasis has been carried out (Yang et al., 2017), a more conclusive meta-analysis of COX-2 in NPC as prognostic biomarker regarding survival outcome was not done before this. Thus, this study aims to determine the prognostic significance of COX-2 in NPC in term of survival outcome using the approach of meta-analysis.

There are a numbers of existing NPC prognosis assessment models such as prognostic nomogram integrating clinical characteristic and hematological/inflammatory biomarkers (Li et al., 2016; Li et al., 2018). Prognostic models to determine death and distant metastasis, disease progression and survival in non-endemic area of China were also reported (Zhang et al., 2013; Zang et al., 2016; Zhao et al., 2019). However, the sample sizes involved in the construction of all of these models were often relatively small. For example, study done by

Zhao and colleagues contained sample size of n=103 only. The patients involved in each of the studies were solely from a particular area and thus, these models only represented and reflected the response on the isolated location of studies. The usage of these models was therefore restricted as these models possessed little/no medical uses in other regions. Additionally, these models focused on clinical characteristic and had little emphasis on the functional genetic prognostic biomarkers. Thus, this study also aims to develop a multigene NPC prognosis assessment model focusing on functional genetic biomarkers by pooling different individual NPC prognostic meta-analysis studies computationally via modification of meta-analysis approach.

Meta-analysis is employed in this study as it provides a more significant and precise result compared to single individual studies. Meta-analysis is a systematic, formal, and quantitative approach and study design used to identify, combine and synthesis the outcomes of previous relevant research. This approach is useful to draw conclusions about a body of studies (Haidich, 2010). The current approach used to conduct meta-analysis is by referring to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher, Liberati, Tetzlaff, Altman & The PRISMA Group, 2009) (refer to Appendix A, page 157-159). By using this method, conflicting results of different individual studies on COX-2 in NPC regarding the OS are pooled and analysed under the same platform to reflect a more significant and relevant output. By modifying PRISMA, different meta-analysis studies focusing on different prognostic factors in NPC are grouped into one platform. This is to integrate isolated aspects of different studies into a more dynamic system and this allows the generation of multigene NPC prognosis assessment model that is based on various functional genetic data. Instead of pooling individual studies that focus on the prognostic significance

of single biomarkers in NPC, pooling of meta-analysis studies of different prognostic biomarkers is therefore more significant and powerful. To our knowledge, this is the first time where modification of PRISMA is applied to generate multigene NPC prognosis assessment model containing different functional genetic biomarkers.

To further understand and explain the multigene NPC prognosis assessment model designed computationally based on data deposited in the databases, a molecular pathway network among the different biomarkers is generated to postulate the physiological interactions among the different biomarkers. This network can clearly display the connections between the different biomarkers in the process of NPC carcinogenesis and hence, explain the complex and dynamic interactions of the pooled biomarkers in the multigene NPC prognosis assessment model.

Herein, the objectives of this project are:

i. To determine the prognostic significance of COX-2 in NPC in term of OS reflected by hazard ratios (HRs) via meta-analysis.
ii. To generate multigene NPC prognosis assessment model focusing on functional genetic biomarkers by pooling different individual NPC prognostic meta-analysis studies via modification of meta-analysis approach.
iii. To explain the multigene NPC prognosis assessment model by generating a molecular pathway network among the different biomarkers.

CHAPTER 2

LITERATURE REVIEW

2.1 Nasopharyngeal Carcinoma (NPC)

2.1.1 Anatomical position of NPC

Nasopharyngeal carcinoma (NPC) is characterized by the malignancy in the tissue of the upper section of pharynx behind the nose (He et al., 2016). NPC is a non-lymphomatous squamous cell carcinoma that particularly occurs in the mucosa epithelial lateral cell lining (including squamous, ciliated pseudostratified columnar and transitional epithelium) (Hong & Kuo, 2003) and the minor salivary glands of the nasopharynx (Sham et al., 1990; Effert et al., 1992; Zong, 2001; Lang, Gao, Guo, Zhao & Zhang, 2014; You et al., 2016). Refer to Figure 2.1, the precise and common site of tumour growth of NPC is the fossa of Rosenmüller, the posteromedial to the medial crura of the Eustachian tube, formed by the intersection of superior and posterior walls of nasopharynx (Sham et al., 1990; Jeyakumar, Brickman & Doerr, 2006; Lee et al., 2012).

2.1.2 Histopatholgy of NPC

According to World Health Organisation (WHO), NPC is histologically subclassified into three subtypes, namely the keratinizing differentiated squamous cell carcinoma (WHO I or Type I), nonkeratinizing carcinoma which can be sub-divided into nonkeratinizing differentiated (WHO II or Type II/IIa) and nonkeratinizing undifferentiated carcinoma (WHO III or Type III/IIb) and basaloid squamous cell carcinoma (Shanmugarantnam & Sobin, 1993; Barnes, Eveson, Reichart & Sidransky, 2005). Nonkeratinizing